



Real-time processing and analytics of sensor data using Kafka and Spark streaming

Postgraduate studies (MSc)

Student: Luke Coetzee (23676469)



Introduction

- Goal: Use streaming databases and predictive analytics to enable real-time results
- Identify and exploit opportunities early / act pre-emptively to circumvent foreseeable obstacles
- Outcome: Artifact creation to leverage real-time analytics and predict outcomes

Key concepts

- Key concept: Streaming database (Kafka)
- Second key concept: Machine learning (Spark)

Research problem

- Businesses have potential to learn from ever-growing amount of data
- Most advantageous form of insight best realized as real-time decisions based on current/applicable events
- Potential for real-time decision making becoming more valuable, if not necessary
- Leverage real-time data analysis offers great advantage

Objectives

- Primary objective:
 - Goal is to leverage real-time analytics by adopting a machine learning-base approach.
 - Recognize early indicators that can be extrapolated to predict outcomes
 - Advantage/disadvantage to a company/institution/user

Objectives (cont.)

- Secondary objectives:
 - Theoretical objectives – a literature review is required for:
 - Streaming databases, specifically Kafka
 - Machine learning
 - Data analytics
 - Empirical objective:
 - Explore the extent to which is viable to gain rapid insight from various sensor data streams by creating an artifact.

Research paradigm and method

- A design science approach is used.
- Main goal: develop knowledge that can be used to design solutions for field problems.
- Understand problem domain by building, and application of designed artifact

Ethical considerations

- NWU ethical process
- Adherence to code of conduct by researcher

Proposed timeline

- The three parts of the proposed timeline, with planned completion end of 2022, are as follows:
 - Formulating – construction and writing
 - Discussions with supervisor to address areas of interests that are missing from section
 - Transformation into the final product

Thank you!

Questions?